TRANSGENIC AMORPHA-4,11-DIENE SYNTHESIS

The present invention relates to a DNA
sequence, a polypeptide encoded by this sequence, and to
the use of said DNA sequence and polypeptide in the
production of amorphadiene.

5   Human malaria is a commonly occurring
widespread infectious disease, caused in 85% of the cases
by <u>Plasmodium falciparum</u>. This parasite is responsible
for the most lethal form of malaria, malaria tropicana.
Each year, malaria causes clinical illness, often very
10 severe, in over 100 million people of which eventually
over 1 million individuals will die. Approximately 40% of
the world's population is at risk of malaria infection
(as estimated by the World Health Organization).

          Malaria has traditionally been treated with
15 quinolines, such as quinine, chloroquine, mefloquine and
primaquine, and with antifolates. Unfortunately, most
<u>P.falciparum</u> strains have become resistant to
chloroquine, and some have developed resistance to
mefloquine and halofantrine as well. Thus, novel
20 antimalarial drugs to which resistant parasites are
sensitive are urgently needed. Artemisinin, as well as
its semisynthetic derivatives are promising candidates
here.

          Artemisinin (Fig. 1), [3R-(3α,5aß,6ß,8aß,
25 9α,12ß,12aR*)]-Octahydro-3,6,9-trimethyl-3,12-epoxy-12H-
pyrano[4,3-j]-1,2-benzodioxepin-10(3H)-one; molecular
weight 282.35), also called arteannuin, qinghaosu or QHS,
is a sesquiterpene lactone endoperoxide isolated from the
aerial parts of the plant <u>Artemisia annua</u> L.

30   <u>Artemisia annua</u> L., also known as quinghao
(Chinese), annual or sweet wormwood, or sweet annie is an
annual herb native to Asia. <u>A.annua</u>, a member of the
Asteraceae, belongs to the tribe Anthemideae of the
Asteroideae, and is a large herb often reaching more than
35 2.0 m in height. It is usually single-stemmed with
alternating branches. The aromatic leaves are deeply
dissected and range from 2.5 to 5 cm in length.
Artemisinin is mainly produced in the leaves as a

secondary metabolite at a concentration of 0.01 - 0.6% on
a dry weight base in natural populations. Artemisinin is
unique to the plant A.annua with one possible exception
of A.apiacea L. The A.annua used in this invention is of
5 Vietnamese origin.

Because of its low concentration in plants,
artemisinin is a relatively expensive resource for a
drug. Current research has thus been aimed at producing
artemisinin at a larger scale by organic synthesis.
10 However, because artemisinin consist of seven chiral
carbon atoms, theoretically $2^7 = 128$ isomers can be formed
of which only one is identical to artemisinin. Because of
this complex structure of artemisinin, production of this
compound by organic synthesis is not profitable from a
15 commercial point of view.

Genetic engineering of the biosynthetic pathway
of artemisinin may give rise to higher artemisinin levels
in plants. To be able to interfere in the biosynthesis of
artemisinin, the biosynthetic pathway has to be known,
20 either completely or partially. Several attempts to
elucidate the entire biosynthetic pathway have been
undertaken. Until now, however, the exact pathway has
remained largely unknown.

In the research that led to the present
25 invention, a unique pathway has been discovered which has
not been published before. This pathway involves inter
alia the formation of the artemisinin precursors amorpha-
4,11-diene (1ß,6ß,7ß,10αH-amorpha-4,11-diene) and the
hydroperoxide of dihydroarteannuic acid. These precursors
30 that were found in A.annua have not been described before
in literature.

From literature it is known that terpene
cyclases (synthases) are branch point enzymes, which
likely play an important role in terpenoid biosynthesis.
35 The working hypothesis for this invention is thus that
over-expression of such a branch point enzyme (terpene
cyclase) may increase terpenoid production in an
organism. Factors that may influence the success of such

an approach are, in the case of artemisinin, the number
and nature of the subsequent biosynthetic steps leading
to artemisinin. Fig. 2 shows the biosynthetic pathway of
artemisinin as postulated by the present inventors.
5 This pathway is divided into three parts:

The first part (Part I) represents the
terpenoid (Isoprenoid) pathway. This pathway is a general
pathway. Farnesyl diphosphate (farnesyl pyrophosphate)
(FPP), for example, is present in every living organism
10 and it is the precursor of a large number of primary and
secondary metabolites. It has been established that FPP
is the precursor of all sesquiterpenes. Thus, by
definition FPP is the precursor of artemisinin.

Part II displays the cyclization of the general
15 precursor FPP into the highly specific precursor amorpha-
4,11-diene (also referred to as amorphadiene), the first
specific precursor of artemisinin. In this pathway
amorphadiene synthase is a branch point enzyme, having a
key position in the biosynthetic pathway of artemisinin.
20   In part III, dihydroarteannuic acid (DHAA),
also called dihydroartemisinic acid, is photo-oxidatively
converted into its hydroperoxide (DHAA-OOH). This
hydroperoxide of DHAA will spontaneously oxidize into
artemisinin. No enzymes are involved in this part of the
25 pathway and therefore it is impossible to alter
artemisinin production by over-expression of genes
involved in this part of the pathway.

Cytochrome P-450 catalyzed enzymes and an
enoate reductase are probably involved in the conversion
30 of amorphadiene into DHAA, the transition state between
part II and part III (see Fig. 3). Because no
intermediates of this part of the pathway are known or
present (accumulated) in detectable amounts, in the
plant, (except arteannuic acid, also called artemisinic
35 acid or 4,11(13)-amorphadien-12-oic acid) it is likely
that these precursors are very rapidly converted into
DHAA. A rate limiting step in this part of the pathway is
not very likely.

Taking al these aspects into account the
inventors concluded that the most logical step to be
altered by genetic interfering, is the conversion
(cyclization) of FPP into amorphadiene by amorphadiene
5 synthase.

The object of the present invention is
therefore to provide a way in which artemisinin can be
obtained via an at least partially biological route.

This object is achieved by the provision of a
10 DNA sequence which exhibits at least a 70% homology to
the sequence as shown in Fig. 12, and which codes for a
polypeptide having the biological activity of the enzyme
amorphadiene synthase.

The biological activity of the enzyme
15 amorphadiene synthase relates to the conversion of the
general precursor farnesyl pyrophosphate (FPP) into the
specific artemisinin precursor amorpha-4,11-diene, which,
in A.annua, is further converted to artemisinin. Suitable
genes according to the invention can be selected by
20 testing the expression product of the gene for its
ability to convert FPP into amorpha-4,11-diene.

By transforming a suitable host cell with the
DNA sequence of the invention, the conversion of farnesyl
pyrophosphate (FPP) into the highly specific precursor
25 amorphadiene can be increased or induced if this
conversion route is not naturally present in the
organism. In the latter case, the organism should
comprise or be able to produce FPP. Suitable host cells
are for example bacterial cells, such as E.coli, yeast
30 cells like Saccharomyces cerevisiae or Pichia pastoris
and in particular oleaginous yeasts, like Yarrowia
lipolytica, or plant cells such as those of A.annua.

Several plants are capable of producing large
amounts of FPP making them potential organisms for
35 amorphadiene production.

The potential oleaginous yeast host cells,
like, for example, Yarrowia lipolytica and Cryptococcus
curvatus, have the capacity to accumulate up to about 50%

(dry weight) of storage carbohydrates in oil bodies,
making them very interesting candidates as production
organisms for large quantities of terpenes. According to
the invention, a way to obtain high levels of terpene
5   accumulation is for example by means of re-direction of
the metabolic flux in favor of the formation of
amorpha-4,11-diene.

In analogy to the approach of an increased
carotenoid production by the food yeast _Candida utilis_
10 through metabolic engineering of the isoprenoid pathway
as done by Shimada et al. (Appl. Environ. Microbiol. **64**,
2676-2680 (1998)) the target genes according to the
invention are acetyl CoA carboxylase (ACC, disruption),
hydroxy-methyl-glutaryl CoA reductase (HMGR,
15 over-expression), and squalene synthase (SQS, disruption)
to obtain an increase of the precursor supplies, and
amorpha-4,11-diene synthase over-expression to obtain
accumulation of amorphadiene in such yeast cells. Because
several expression systems (for example Muller et al.,
20 Yeast **14**, 1267-1283 (1998); Park et al., The Journal of
Biological Chemistry **272**, 6876-6881 (1997); Tharaud et
al., Gene **121**, 111-119 (1992)) and transformation systems
(for example Chen et al., Appl. Microbiol. Biotechnol.
48, 232-235 (1997)) are known for _Y.lipolytica_ in
25 literature, transformation and expression of the
previously mentioned target genes in _Y.lipolytica_ is
possible without serious technical problems.

By adding FPP to a culture medium further
comprising the enzyme of the invention (isolated as
30 described in example 1), or transformed cells, e.g.
E.coli, comprising the DNA sequence of the invention (as
described in examples 3 and 4), which is expressed, FPP
is converted into amorphadiene. Amorphadiene can then be
used as a starting material for the production of
35 artemisinin.

Transformed cells in which amorphadiene is
produced as a result of the expression of amorphadiene
synthase of the invention can be used either in disrupted

form, by for example sonication, or as intact cells, as a source of amorphadiene.

Over-expression of the amorphadiene synthase encoding gene in A.annua will increase artemisinin
5   production, because the terpene cyclase is expected to be the rate limiting step.

The results of the present research (postulated biosynthetic pathway of artemisinin) make the presence of a single major rate limiting step at the place of the
10 amorphadiene synthase clear. Over-expression of the amorphadiene synthase encoding gene can increase the production of artemisinin in A.annua.

The chemical structure of the first specific precursor of artemisinin, a cyclization product of FPP,
15 was not known in literature. Neither has anyone so far detected such a compound in A.annua. Nevertheless it was possible to predict a likely structure for this cyclization product, based on the structure of DHAA and arteannuic acid (Fig. 3). The structure predicted in this
20 way was consistent with a compound which is known in literature as 4,11-amorphadiene (J.D. Connelly & R.A. Hill in: Dictionary of terpenoids, Chapmann and Hill, London, England), as depicted in Fig. 4. This compound, isolated from Viguiera oblongifolia, has previously been
25 described by Bohlmann et al. under the incorrect name cadina-4,11-diene (Phytochemistry 23(5) 1183-1184 (1984)). Starting from arteannuic acid (isolated from A.annua), it was possible to synthesize amorphadiene. Amorphadiene obtained in this way was in all chemical and
30 physical aspects identical to amorphadiene as described by Bohlmann et al., and this standard was used to show the presence of amorphadiene in a terpene extract of A.annua.

A further object of the present invention is to
35 provide a polypeptide having the biological activity of the enzyme amorphadiene synthase, obtainable by a process as described in example 1. This polypeptide can be used to convert FPP into amorphadiene which subsequently can

be converted into artemisinin. Conversion can take place
either _in planta_, when the polypeptide amorphadiene
synthase is expressed in a plant that contains the
necessary enzymes to further convert amorphadiene into
5 artemisinin, or _in vitro_ when FPP and the polypeptide
(either in isolated form or as an expression product in a
cell) are brought together in an incubation mixture.

Amorphadiene, produced by a suitable host
organism transformed with the DNA sequence of the
10 invention as precursor, can subsequently be chemically
converted to dihydroarteannuic acid. Dihydroarteannuic
acid _per se_ can be used or in the production of
artemisinin.

The chemical conversion of amorphadiene into
15 dihydroarteannuic acid (Fig. 15) starts with the
enantio-, stereo- and regioselective (anti-markownikoff)
hydroboration of amorphadiene with BH_3, yielding a
trialkylborane, followed by an oxidation of the
trialkylborane with $NaOH/H_2O_2$ yielding the alcohol
20 (Advanced Organic Chemistry, Jerry March, 4th Edition,
Wiley, 1992). A mild oxidation of the alcohol to the acid
can be obtained by PDC (pyridinium dichromate) without
attacking the second double bond (Fig. 15) (Organic
Synthesis, M.B. Smith, 1st Edition, McGraw-Hill, 1994).
25   Many genes encoding enzymes involved in the
biosynthetic pathway of farnesyl diphosphate are cloned
and known in literature. For _A.annua_, for example, the
sequence of the farnesyl diphosphate synthase encoding
gene is known in literature (Y. Matsushita, W-K. Kang and
30 V. Charlwood Gene, 172 (1996) 207-209). A further
approach to introduce or increase the amorphadiene
production in an organism, is to transform such an
organism (for example _A.annua_) simultaneously with the
DNA sequence of the invention with one or more genes
35 involved in the biosynthesis of farnesyl diphosphate. The
expression of a fusion protein of amorphadiene synthase
and farnesyl diphosphate synthase may be an example here.

(Sesqui)terpenes, such as amorphadiene, are
also known as flavor and fragrance compounds in the food
and perfume industry. In addition, terpenes play a role
in plant-insect interactions, such as the attraction or
5 repulsion of insects by plants. Furthermore, dihydro-
arteannuic acid, which is an intermediate in the
metabolic route from amorphadiene into artemisinin in
A.annua, can be used as an antioxidant.

Amorphadiene, obtained by (over)expression of
10 the DNA sequence of the invention, or by using the
polypeptide (amorphadiene synthase) of the invention, can
be applied for these purposes as well.

The plants that can be used for this invention
are preferably plants already producing artemisinin. A
15 prime example is Artemisia annua, as this species
contains the remainder of the pathway leading to
artemisinin. However, this invention may also be used for
the production of amorphadiene in plants, which, as
mentioned before, can be used as a flavor or fragrance
20 compound or biocide, or can be converted to artemisinin,
either chemically or by bioconversion using
microorganisms, yeasts or plant cells.

The plant that can be used for the production
of amorphadiene is preferably a plant already producing
25 sesquiterpenes, as these plants already have the basic
pathway and storage compartments available, or a plant in
which the biosynthesis of sesquiterpenoids can be induced
by elicitation. The methods of this invention are readily
applicable via conventional techniques to numerous plant
30 species, including for example species from the genera
Carum, Cichorium, Daucus, Juniperus, Chamomilla, Lactuca,
Pogostemon and Vetiveria, and species of the inducible
(by elicitation) sesquiterpenoid phytoalexin producing
genera Capsicum, Gossypium, Lycopersicon, Nicotiana,
35 Phleum, Solanum and Ulmus. However, also common
agricultural crops like soybean, sunflower and rapeseed
are interesting candidates here.

9

The invention will be further illustrated by
the following examples, but will not be limited thereto.
In the examples reference is made to the following
figures:

5   **Fig. 1**: Structural formula of artemisinin.

              Fig. 2: Postulated biosynthetic pathway of
artemisinin in A.annua.

              Fig. 3: Transition between part II and III of
Fig. 2: hypothetical conversion of amorphadiene into
10 dihydroarteannuic acid in A.annua.

              Fig. 4: Structural formula of amorpha-4,11-
diene.

              Fig. 5: Radio-GC chromatograms of the [^{3}H]-FPP-
assays. A. Flame Ionization Detector (FID) signal of
15 amorphadiene (reference). B. Radio signals of the ^{3}H
labeled assay products amorphadiene (retention time 14
min.) and farnesol (as a product of aspecific
phosphohydrolase activity, retention time 28 min.)
obtained with crude enzyme extract. C. Radio signal of
20 the ^{3}H labeled assay product amorphadiene obtained with
Mono-Q purified enzyme extract.

              Fig. 6: Mass spectrum of reference amorphadiene
compared with mass spectrum of the FPP assay with terpene
cyclases (synthases) purified from A.annua. This
25 comparison yielded a quality score of 99%, corresponding
with a maximum score of identicalness.

              Fig. 7: Probe generated by PCR and cloned into
pGEM 7Zf⁺.

              Fig. 8: Nucleotide sequence and deduced amino
30 acid sequence of the probe (538 bp) generated by PCR with
primers A and B.

              Fig. 9: Released plasmid of a positive clone
isolated from the cDNA library of induced A.annua.

              Fig. 10: Nucleotide sequence and deduced amino
35 acid sequence of a positive clone (amorphadiene synthase
encoding gene) isolated from the cDNA library of induced
A.annua. The sequence is flanked with EcoRI (NotI)
adapters (Gibco BRL).

Fig. 11: Part, between start and stop codon (flanked by <u>Nco</u>I and <u>Bam</u>HI sites, respectively), of the amorphadiene synthase encoding gene cloned in the NcoI/<u>Bam</u>HI site of the expression vector pET 11d.

5   **Fig. 12**: Nucleotide sequence and deduced amino acid sequence of the amorphadiene synthase encoding gene, between start and stop codon (flanked by <u>Nco</u>I and <u>Bam</u>HI sites, respectively), obtained by PCR with primers C and D.

10   **Fig. 13**: SDS-PAGE gel: lanes 1 and 2 show pellet and supernatant of pET 11d, respectively (negative control); lanes 3 and 4 show pellet and supernatant of tobacco 5-<u>epi</u>-aristolochene synthase (TEAS) gene in pET 11d (positive control), lanes 5, 7, 9 and 6, 8, 10,
15 respectively show pellet and supernatant of amorphadiene synthase in pET 11d. All constructs were expressed in <u>E.coli</u> BL21 (DE3). The lanes with the pellet fractions of TEAS in pET 11d (positive controls) and amorphadiene synthase in pET 11d show a clear spot which was not
20 present in the negative control pET 11d. Mw is low Molecular Weight marker (Pharmacia Biotech).

          Fig. 14: A. Flame Ionization Detector (FID) signals of amorpha-4,11-diene and farnesol (references); B. Radio-GC chromatograms of the [³H]-FPP-assays with
25 intact BL21 (DE3) cells, transformed with the amorphadiene synthase encoding gene in the expression vector pET 11d; C. Radio-GC chromatograms of the [³H]-FPP-assays with the supernatant of sonicated BL21 (DE3) cells, transformed with the amorphadiene synthase
30 encoding gene in the expression vector pET 11d.

          Fig. 15: Hypothetical chemical synthesis of dihydroarteannuic acid using amorpha-4,11-diene as a precursor. The reaction consists of an enantio-, stereo- and region selective (anti-markownikoff) hydroboration of
35 amorphadiene with BH₃ followed by an oxidation of the formed trialkylboranes with NaOH/H₂O₂ yielding the alcohol. A mild oxidation of the alcohol to the acid can

be obtained with PDC (pyridinium dichromate) without
attacking the second double bond.

Fig. 16: Determination of the molecular weight
of amorpha-4,11-diene synthase by size-exclusion
5 chromatography (gel filtration). -*- is activity curve;
-▲- is molecular weight markers; —— is molecular weight
calibration line.

10 **EXAMPLES**

EXAMPLE 1

Conversion of farnesyl pyrophosphate into amorphadiene by
amorphadiene synthase

A. Isolation, partial purification and identification of
15 amorphadiene synthase from A.annua

During enzyme isolation and preparation of the
assays, all operations were carried out on ice or at 4°C.
Ten grams of frozen young leaves from greenhouse-grown
A.annua were ground in a pre-chilled mortar and pestle in
20 40 ml of pre-chilled buffer containing 25 mM MES (pH
5.5), 20% (v/v) glycerol, 25 mM sodium ascorbate, 25 mM
NaHSO$_3$, 10 mM MgCl$_2$ and 5 mM DTT (buffer A) slurried with
1 g polyvinylpolypyrrolidone (PVPP) and a spatula tip of
purified sea sand. Ten grams of polystyrene resin
25 (Amberlite XAD-4, Serva) were added and the slurry was
stirred carefully for 10 min and then filtered through
cheesecloth. The filtrate was centrifuged at 20,000g for
20 min (pellet discarded), and then at 100,000g for 90
min. A 3-ml subsample of the supernatant was desalted to
30 a buffer containing 15 mM MOPSO (pH 7.0), 10% (v/v)
glycerol, 1 mM sodium ascorbate, 10 mM MgCl$_2$ and 2 mM DTT
(buffer B) and used for enzyme assays/product
identification (see below at 'B').

The remainder of the supernatant was added to
35 12.5 g DEAE anion exchanger (Whatman DE-52), which had
been rinsed several times with buffer A, and stirred
carefully for 10 min. After centrifugation at 18,000g for
20 min, the supernatant was decanted and the DE-52 pellet

12

discarded. Proteins in the supernatant were precipitated
by adding $(NH_4)_2SO_4$ to a final concentration of 70%,
careful stirring for 30 min, and centrifugation at
20,000g for 10 min. The resulting pellet was resuspended
5 in 6 ml buffer A and desalted to buffer B. After addition
of glycerol up to 30% (v/v) this enzyme preparation could
be frozen in liquid N_2 and stored at -80°C without loss of
activity. 0.5 ml of this enzyme preparation was applied
to a Mono-Q FPLC column (HR5/5, Pharmacia Biotech),
10 previously equilibrated with buffer B without sodium
ascorbate, with 0.1% Tween-20. The enzyme was eluted with
a gradient of 0-2.0 M KCl in the same buffer. For
determination of enzyme activities, 50 μl of the 0.75-ml
fractions were diluted 2-fold in an Eppendorf tube with
15 buffer B and 20 μM [^{3}H]FPP was added. The reaction mixture
was overlaid with 1 ml of hexane to trap volatile
products and the contents mixed. After incubation for 30
min at 30°C, the vials were vigorously mixed, and
centrifuged briefly to separate phases. A portion of the
20 hexane phase (750 μl) was transferred to a new Eppendorf
tube containing 40 mg of silica gel (0.035-0.07 mm, pore
diameter 6 nm, Janssen Chimica) to bind terpenols
produced by phosphohydrolases, and, after mixing and
centrifugation, 500 μl of the hexane layer was removed
25 for liquid scintillation counting in 4.5 ml of Ultima
Gold cocktail (Packard). The active fractions were
combined, and an assay carried out to determine product
identity (see below). After the Mono-Q step, the enzyme
was separated from all other FPP-converting activities
30 (Fig. 5C). This enzyme preparation was used for the
measurement of enzyme characteristics such as molecular
weight and K_m. The molecular weight was determined using
size-exclusion chromatography. 200 μl of the Mono-Q
eluent was loaded on a Superdex 75 (H/R10/30, Pharmacia
35 Biotech) and eluted in the same buffer as used for Mono-
Q. Enzyme activities in 0.5 ml fractions were determined
as described for Mono-Q, but using undiluted eluent. The
column was calibrated using cytochrome C, ribonuclease A,

α-chymotrypsinogen, ovalbumin and BSA (all from Sigma).
The estimated molecular weight was 56 kDa (Fig. 16).
Enzyme-kinetics were determined using 5- and 10-fold
diluted Mono-Q eluted enzyme preparation and [^{3}H]-FPP
5 concentrations ranging from 0.25-100μM. K_m for
amorphadiene synthase was 0.6 μM.

B. Determination of product identity
        For determination of product identity, 20 μM
10 [^{3}H]-FPP (Amersham; for radio-GC analysis) or 50 μM
unlabelled FPP (Sigma; for GC-MS analysis) were added to
1 ml of the enzyme preparations. After the addition of a
1 ml redistilled pentane overlay to trap volatile
products, the tubes were carefully mixed and incubated
15 for 1 h at 30°C. Boiled samples were used as controls.
Following the assay, the tubes were vigorously mixed. The
organic layer was removed and passed over a short column
of aluminum oxide overlaid with anhydrous $MgSO_4$. The assay
was extracted with another 1 ml of diethyl ether which
20 was also passed over the aluminum oxide column, and the
column washed with 1.5 ml of diethyl-ether. For GC-
analysis, the combined pentane/diethyl-ether mixture was
slowly concentrated under a stream of N_2.
        Radio-GLC was performed on a Carlo-Erba 4160
25 Series gas chromatograph equipped with a RAGA-90
radioactivity detector (Raytest, Straubenhardt, Germany).
Sample components eluting from the column were
quantitatively reduced before radioactivity measurement
by passage through a conversion reactor filled with
30 platinum chips at 800°C. Samples of 1 μl were injected in
the cold on-column mode. The column was a fused silica
capillary (30 m x 0.32 mm i.d.) coated with a film of
0.25 μm of polyethylene glycol (EconoCap EC-WAX, Alltech
Associates) and operated with a He-flow of 1.2 ml min^{-1}.
35 The oven temperature was programmed to 70°C for 5 min,
followed by a ramp of 5° min^{-1} to 210°C and a final time
of 5 min. To determine retention times and peak
identities (by co-elution of radioactivity with reference

standards), about 20% of the column effluent was split
with an adjustable splitter to an FID (temperature
270°C). The remainder was directed to the conversion
reactor and radio detector. H_2 was added prior to the
5   reactor at 3 ml min^{-1}, and CH_4 as a quench gas prior to the
radioactivity detector (5 ml counting tube) to give a
total flow of 36 ml min^{-1}. The major [^{3}H]-labeled product
co-eluted with the amorphadiene reference standard
(retention time 14 min) (Fig. 5B). The second
10 radiolabeled product is farnesol, the product of
aspecific phosphohydrolase activity. After the Mono-Q
step, the enzyme was separated from all other FPP-
converting activities (Fig. 5C). This enzyme preparation
was used for the measurement of enzyme characteristics
15 such as molecular weight and K_m.

        GC-MS analysis was performed using a HP 5890
series II GC and HP 5972A Mass Selective Detector
(Hewlett-Packard) equipped with an HP-5MS or HP-Innowax
column (both 30 m x 0.25 mm i.d., 0.25 μm df). The oven
20 was programmed at an initial temperature of 70°C for 1
min, with a ramp of 5°C min^{-1} to 210°C and final time of 5
min. The injection port (splitless mode), interface and
MS source temperatures were 175, 290 and 180°C,
respectively, and the He inlet pressure was controlled by
25 electronic pressure control to achieve a constant column
flow of 1.0 ml min^{-1}. Ionization potential was set at 70
eV, and scanning was performed from 30-250 amu. The
$(NH_4)_2SO_4$ precipitated enzyme preparation was free of
endogenous sesquiterpenes. GC-MS analysis on the two
30 different GC-columns of sesquiterpene products generated
from FPP by this enzyme preparation showed that the main
product had a mass spectrum and retention time equal to
that of the semi-synthetically produced amorphadiene
(Fig. 6).

35

EXAMPLE 2
Isolation and characterization of the amorphadiene
synthase encoding gene
A. Induction of transcription

5   As revealed in part III of Fig. 2, DHAA is
photo-oxidatively converted into DHAA-OOH. In this
reaction a reactive form of oxygen (singlet O_2) is added
to DHAA. DHAA plays the role of an anti-oxidant, a
scavenger of reactive oxygen species. Artemisinin is the

10 stable end product of this reaction in which reactive
oxygen is stored. Under stress conditions, (for example
photo-stress, frost, drought or mechanical damage)
reactive species of oxygen are formed in the plant. In
response to this reactive oxygen generally plants are

15 producing anti-oxidants. It is likely that A.annua will
produce DHAA as anti-oxidant in response to this release
of reactive oxygen. By exposing A.annua to stress
conditions the transcription of the gene encoding
amorphadiene synthase will be induced. To achieve this

20 situation A.annua plants grown under climate room
conditions (23°C, 90% moisture, 3000 lux) were exposed to
stress conditions by putting them for one hour at
approximately 30% moisture (drought stress) and 6000 lux
(photo stress) at 30°C.

25

B. Isolation of total RNA
        Total RNA of stress induced plants (according
to example 2.A) was isolated from young leaves by the
method of Verwoerd et al. (Nucleic Acids Research **17**(6),

30 2362 (1989)). DNase I (Deoxyribonuclease I, RNase free)
was used to remove DNA from the RNA isolate. The DNase I
was inactivated by exposure at 70°C during 15 minutes.

C. cDNA synthesis

35   The reverse transcription reaction was carried
out in a 20 μl reaction containing 5 μg total RNA, 0.2 μg
oligo (dT)$_{12}$, 0.5 mM each dATP, dTTP, dCTP and dGTP, 10 mM
DTT, 2 U ribonuclease inhibitor (Gibco BRL), first strand

synthesis buffer (Promega) and catalyzed with 200 U
Moloney murine leukemia virus (M-MLV) reverse
transcriptase RNase H minus (Promega). After 1 h
incubation at 37°C the reaction was stopped by storing
5   the reaction mixture at -20°C.

D. PCR-based probe generation

        Based on comparison of sequences of terpenoid
synthases, two degenerated primers were designed for two
10   conserved regions. The sequence of the sense primer
(primer A) was 5'-GA(C/T) GA(G/A) AA(C/T) GGI AA(G/A)
TT(C/T) AA(G/A) GA-3' and the sequence of the anti sense
primer (primer B) was 5'-CC (G/A)TA IGC (G/A)TC (G/A)AA
IGT (G/A)TC (G/A)TC-3'. PCR was performed in a total
15   volume of 100 μl containing 0.5 μM of each of these two
primers, 0.2 mM each dNTP, 1 U Super Taq polymerase / 1x
PCR buffer (HT Biotechnology LTD, Cambridge, England) and
2 μl cDNA. The reaction was incubated in a thermocycler
(PTC 150, MJ-research) with 1 minute denaturation at
20   95°C, 1 minute annealing at 40°C and 1 minute and 15
seconds elongation at 72°C during 40 cycles. Agarose gel
electrophoresis revealed a single specific PCR product of
approximately 550 bp (538 bp). Such a specific
amplification product was only obtained when using cDNA
25   made of RNA isolated from stress induced plants. The PCR
product was made blunt by using DNA polymerase I large
fragment (Klenow), gel-purified and subcloned in Sma I
digested pGEM 7Zf(+) (Stratagene) (Fig. 7) and E.coli
DH5α (Gibco BRL) was transformed with this construct. The
30   inserts of 8 individual transformants were sequenced and
they all had the same sequence as shown in Fig. 8.

E. cDNA Library construction

        Synthesis of the second strand of the cDNA was
35   done analogous to the RiboClone® cDNA synthesis System
(Promega). After ligation with EcoR I (Not I) adapters
(Gibco BRL) with sequence:

```
5'-pGTCGACGCGGCCGCG-3'
3'-CAGCTGCGCCGGCGCTTAA-OH-5'
```

the double stranded DNA was ligated into λ ExCell

5   EcoRI/CIP (Pharmacia Biotech). For packaging and plating
of the cDNA library, the Ready-To-Go® Lambda Packaging
Kit (Pharmacia Biotech) was used. The titer of the
unamplified library was 1.2 x 10^6 plaque forming units.

10 F. Library screening

For library screening 200 ng of the PCR
amplified probe (Fig. 8) was gel purified, randomly
labeled with [α-^{32}P]dCTP, according to the manufacturer's
recommendation (Random Primed DNA Labeling Kit,

15 Boehringer Mannheim Biochemica) and used to screen
replica filters of 10^4 plaques of the cDNA library plated
on E.coli NM 522. The hybridization was performed for 16
h at 68°C in 1 M NaCl, 1% SDS and 10% PEG (5000-7000).
Filters were washed two times for 10 minutes at 50°C in 2

20 x SSC with 0.1% SDS and exposed for 16 h to a Fuji X-ray
film at -70°C. Clones yielding positive signals were
isolated through a second and third round of
hybridization. By transfecting E.coli NP66 (Pharmacia
Biotech) with the positive clones, plasmid releases (Fig.

25 9) were obtained according to the manufacturer's
instructions (Pharmacia Biotech). Sequencing of these
positive clones yielded a sequence as revealed in Fig.
10.

30 **EXAMPLE 3**

Expression of the amorphadiene synthase encoding gene in
E.coli BL21 (DE3)

For functional expression the cDNA clone was
subcloned in frame into the expression vector pET 11d

35 (Stratagene). To introduce suitable restriction sites for
subcloning, the gene was amplified by PCR using a sense
primer (primer C) 5'-GTCGACAAACCATGGCACTTACAGAA G-3'

(introducing a NcoI site at the start codon ATG) and an anti-sense primer (primer D):
5'-GGATGGATCCTCATATACTCATAGGATAAACG-3' (introducing a BamHI site directly behind the stop codon TGA). The PCR
5   reaction was performed under standard conditions. After digestion with BamHI and NcoI, the PCR product (Fig. 12) and the expression vector pET 11d were gel purified and ligated together to yield a construct as revealed in Fig. 11.

10   To obtain expression, this gene construct (Fig. 11), pET 11d without an insert as negative control, and pET 11d with the tobacco 5-epi-aristolochene synthase (TEAS) gene (Back et al., Archives of Biochemistry and Biophysics 315(2) 527-532 (1994); Facchini & Chappell,
15   Proc. Natl. Acad. Sci. USA 89, 11088-11092 (1992); Back & Chappell, The Journal of Biological Chemistry 270, 7375-7381 (1995)) as positive control were transformed to E.coli BL21(DE3) (Stratagene), and grown overnight on LB agar plates supplemented with ampicillin at 37°C.
20   Cultures of 50 ml LB medium supplemented with ampicillin (100 μg/ml) and 0.25 mM isopropyl-1-thio-ß-D-galactopyranoside (IPTG) were inoculated with these over night cultures to A_{600} = 0.5 and grown for 3 h at 27°C. The cells were harvested by centrifugation during 8 minutes
25   at 2000 g and resuspended in 2 ml assay buffer. An aliquot of 1 ml resuspended cells was sonicated on ice four times for 5 seconds with 30 second intervals, centrifuged for 5 minutes at 4°C in a microfuge (13.000 rpm) and the supernatant used for cyclase enzyme activity
30   determinations and SDS-PAGE gel electrophoresis.
             Expression of the amorphadiene synthase gene-pET 11d construct (Fig. 11) in E.coli BL21(DE3) yielded a protein of approximately 50 to 60 kDa as shown in Fig. 13 lane 5 to 10. This agrees well to the size of
35   amorphadiene synthase isolated from A.annua, which was determined to be 56 kDa (Fig. 16).

EXAMPLE 4

Conversion of FPP into amorphadiene by amorphadiene
synthase expressed in E.coli.

Besides the supernatant of sonicated cells,

5 also intact cells were used in the FPP assay. The FPP
assay, GC-RAGA and GC-MS analyses were performed as
described previously. Figs. 14 and 14A are revealing the
GC-RAGA chromatograms of the assays with intact
transformed cells and with the supernatant of sonicated

10 transformed cells, respectively. In both assays
amorphadiene was produced. Identification of these assay
products with the GC-MS gave a mass-spectrum identical to
the mass-spectrum of the reference amorphadiene with a
quality score of 99% (maximum score), mass spectra were

15 identical to the spectra as shown in Fig. 6. No
amorphadiene was found in assays done with the positive
and negative controls.

EXAMPLE 5

20 Expression of amorpha-4,11-diene synthase in transgenic
tobacco

There are many ways to introduce DNA into a
plant cell. Suitable methods by which DNA can be
introduced into the plant cell include Agrobacterium

25 infection or direct delivery of DNA such as, for example,
by PEG-mediated transformation of protoplasts (Omirulleh
et al., Plant Molecular Biology **21**, 415-428 (1993)) or
electroporation, by acceleration of DNA coated
microprojectiles (for example, microprojectile

30 bombardment) microinjection, etc.

Because Agrobacterium tumefaciens-mediated
transformation of Artemisia annua and Nicotiana tabacum
with a sesquiterpene cyclase gene is known in literature
(Vergauwe et al., Plant Cell Reports **15**, 929-933 (1996);

35 Hohn and Ohlrogge, Plant Physiol. **97**, 460-462 (1991))
delivery of expression units (cassettes), containing the
amorphadiene synthase encoding gene, mediated by
Agrobacterium seemed to be a rational approach.

There are several binary vector systems
suitable to transfer the amorphadiene synthase encoding
gene assembled in an expression cassette behind a
suitable promoter (for example, the cauliflower mosaic
5 virus (CaMV) 35S promoter) and upstream of a suitable
terminator (for example, the nopaline synthase
transcription terminator (nos-tail)) to tobacco and/or
A.annua.

        Analogous to EXAMPLE 3, suitable restriction
10 sites for subcloning were introduced by using PCR with a
sense primer (primer G) 5'-GA GGA TCC **ATG** TCA CTT ACA
GAA-3' introducing a BamHI site preceding the start codon
ATG) and an anti-sense primer (primer H) 5'-AT GGA TCC
TCA TAT ACT CAT AGG A-3' (introducing a BamHI site
15 directly behind the stop codon TGA). After digestion with
BamHI the PCR product and the plant-expression cassette
pLV399 were gel purified and ligated to provide the gene
encoding amorpha-4,11-diene synthase with the cauliflower
mosaic virus 35S promoter and a nopaline synthase
20 transcription terminator. The plant-expression casette
pLV399 is a pUC 19 vector (Yanisch-Perron, C. et al.,
Gene 33, 103-119 (1985)) in which the multiple cloning
site (polylinker) is replaced by a CaMV 35 S promoter
BamHI fused to a nos-tail (terminator) flanked by the
25 'unique' sites; EcoRI, KpnI, XhoI, and a HindIII site
downstream from the promoter and EcoRI, XhoI, PstI, SphI,
KpnI, HindIII upstream from the terminator. The
orientation of the amorpha-4,11-diene encoding gene in
pLV399 was checked by restriction analysis with PstI and
30 NdeI. After partial digestion of this construct with KpnI
the amorpha-4,11-diene encoding gene flanked by the 35S
promotor and nos terminator was ligated into the KpnI
digested binary vector pCGN1548.

        To mobilize the recombinant binary vector to
35 Agrobacterium tumefaciens LBA4404 (Gibco BRL, Life
Technologies), a triparental mating procedure was carried
out by using E.coli (DH5α) carrying the recombinant
binary vector and a helper E.coli carrying the plasmid

pRK2013 to mobilize the recombinant binary vector to A. tumefaciens LBA4404.

This transformed Agrobacterium strain was used for transformation of explants from the target plant
5 species. Only the transformed tissue carrying a resistance marker (kanamycin-resistance, present between the binary plasmid T-DNA borders) regenerated on a selectable (kanamycin containing) regeneration medium. (According to Rogers SG, Horsch RB, Fraley RT Methods
10 Enzymol (1986)118: 627-640).

The plants regenerated out of the transformed tissue expressed the amorphadiene synthase gene as followed from the presence therein of amorphadiene as confirmed by GC-MS analyses.
15

EXAMPLE 6

Conversion of amorphadiene into artemisinin (DHAA) by A.annua

This assay was carried out in a way analogous
20 to the method as described by Koepp et al. (The Journal of Biological Chemistry **270**, 8686-8690 (1995)). Radioactive (^{3}H-labeled) amorphadiene was fed to leaf discs of A.annua. For the infiltration of amorphadiene into the leaf discs of A.annua the radioactive
25 amorphadiene can be made water soluble by complexation with cyclodextrins, for example. Radioactive amorphadiene is obtained by using the FPP-assay with the transformed E.coli BL21(DE3) cells (carrying the cloned amorphadiene synthetase gene of A.annua). Identification of the
30 product(s) made in this assay was done by radio-GC analysis. The expected intermediates arteannuic acid (AA), dihydroarteannuic acid (DHAA) and the end product artemisinin were all used as references.

A mixture of α-cyclodextrin, ß-cyclodextrin, γ-
35 cyclodextrin, and partially ^{3}H-labeled amorpha-4,11-diene (20 μM) in a molar ratio of 5:5:5:1 was prepared and A.annua leaf discs were incubated in this mixture. After 120 hours of incubation artemisinic acid and

22

dihydroartemisinic acid could be detected by radio-GC in
a way analogous to part B of example 1.

EXAMPLE 7

5 <u>Expression of amorpha-4,11-diene synthase in transgenic
A.annua and the production of artemisinin</u>

Transformed <u>A.annua</u> plants were prepared as
described in example 5.

For the regeneration of <u>A.annua</u> the medium for
10 callus, shoot and root induction consisted of Murashige
and Skoog micro and macro elements including modified
vitamins (Duchefa Biochemie, Haarlem, The Netherlands),
4% (w/v) sucrose, 0.1 mg/L Indole-3-acetic acid (IAA),
0.1 mg/L 6-benzylaminopurine (BAP) and 0.8% (w/v) agar
15 (Plant agar, Duchefa Biochemie, Haarlem, the
Netherlands). The pH was adjusted to 5.7 with NaOH prior
to the addition of agar. The medium was autoclaved at 1
bar for 20 min. Transformed explants were regenerated on
this medium to fully regenerated plants.
20   The regenerated plants were found to over-
express the enzyme amorpha-4,11-diene synthase which led
to production of artemisinic acid, dihydroartemisinic
acid, and artemisinin at a level above the natural level
in non-transformed plants.
25

EXAMPLE 8

<u>Expression of the amorpha-4,11-diene synthase gene in
Saccharomyces cerevisiae and Pichia pastoris</u>

For functional expression the cDNA clone was
30 subcloned into the inducible expression vector pYES2
(episomal vector, Invitrogen) and the constitutive
expression vector (integrating the gene construct into
the genome) pGAPZ A (Invitrogen). To introduce suitable
restriction sites for subcloning, the gene was amplified
35 by PCR using a sense primer (primer E) 5'-CGA <u>GAA TTC</u> **ATG**
TCA CTT ACA G-3' (introducing a <u>EcoRI</u> site preceding the
start codon **ATG**) and an anti-sense primer (primer F)
5'-GGAT <u>CTC GAG</u> **TCA** TAT ACT CAT-3' (introducing a <u>BamHI</u>

site directly behind the stop codon **TGA**). Subcloning of
the PCR product into pYES2 and pGAPZ A was done in a way
analogue to Example 3.

5   The obtained gene constructs were transformed
to respectively <u>Saccharomyces cerevisiae</u> and <u>Pichia
pastoris</u> using the <u>S.cerevisiae</u> EasyComp™ transformation
kit (Invitrogen) to transform <u>S.cerevisiae</u> and the <u>Pichia</u>
EasyComp™ transformation kit (Invitrogen) for
transformation of <u>P.pastoris</u>. All transformations were
10 carried out according to the instructions of the
manufacturer. Growth, selection and induction were also
performed in accordance to the instructions of the
manufacturer. Harvesting and sonication of the yeast
cells was done in an analogous way to the method as
15 described in Example 3.

The FPP assay with the extracts of the yeast
cells in which the amorpha-4,11-diene synthase gene was
expressed yielded identical GC-RAGA and GC-MS
chromatograms as obtained in example 4.

20